

# Semi-Supervised Segmentation based on Non-local Continuous Min-Cut

Nawal Houhou<sup>1</sup>, Xavier Bresson<sup>2</sup>, Arthur Szlam<sup>2</sup>,  
Tony F. Chan<sup>2</sup>, and Jean-Philippe Thiran<sup>1</sup>

<sup>1</sup> Signal Processing Laboratory 5, Swiss Federal Institute of Technology (EPFL),  
Lausanne, Switzerland

<sup>2</sup> Department of Mathematics, University of California  
Los Angeles, CA 90095-1555, USA

**Abstract.** We propose a semi-supervised image segmentation method that relies on a non-local continuous version of the min-cut algorithm and labels or seeds provided by a user. The segmentation process is performed via energy minimization. The proposed energy is composed of three terms. The first term defines labels or seed points assigned to objects that the user wants to identify and the background. The second term carries out the diffusion of object and background labels and stops the diffusion when the interface between the object and the background is reached. The diffusion process is performed on a graph defined from image intensity patches. The graph of intensity patches is known to better deal with textures because this graph uses semi-local and non-local image information. The last term is the standard TV term that regularizes the geometry of the interface. We introduce an iterative scheme that provides a unique minimizer. Promising results are presented on synthetic textures and real-world images.

## 1 Introduction

Image segmentation is an important problem in image processing. The objective of segmentation algorithms is to partition an image into a finite number of semantically important regions such as anatomical or functional structures in medical images or objects in natural images. Well-posed approaches to solve the image segmentation problem are energy minimization methods. This paper introduces an energy minimization algorithm to solve the semi-supervised segmentation problem based on the continuous min cut/max flow model originally defined by Strang in [1]. Semi-supervised segmentation models defined in a continuous setting have already been proposed in the literature. Among them, Protiere and Sapiro proposed in [2] an interactive algorithm for segmentation. Cremers et al. introduced in [3] an algorithm based on the level set method to perform interactive image segmentation. Appleton and Talbot introduced in [4] a semi-supervised segmentation model based on the continuous min-cut of Strang. Unger et al. defined in [5] a segmentation method also based on the min-cut model of Strang in [1]. The semi-supervised segmentation models using

the continuous min-cut are based on local image information. These models s.a. [4, 5] perform very well for the segmentation of smooth regions, but they are less efficient with textures. In this paper, we extend the continuous min cut to a non-local formulation along the same line as non-local means defined by Buades, Coll and Morel in [6] and the variational non-local means model of Gilboa and Osher [7]. This non-local extension of the continuous min cut can be obtained in different ways. We used the original discrete min-cut model [8] to define the non-local continuous min-cut, which turns to be the  $H^1$  norm defined on graph and a term that constraints the labels. The  $H^1$  norm carries out the diffusion of object and background labels on the graph of image patches [9, 6], which holds semi-local and non-local image information that can better segment textures and real-world objects. Besides, the continuous formulation of the min-cut algorithm allows us introducing other regularization processes such as the TV energy. The TV energy is indeed useful to regularize the boundary between the object and the background. Besides, the TV energy can smooth out the segmentation of small sets favored by the min-cut algorithm as noticed by Shi and Malik in [10] (see Figure 1).

## 2 Graph, Min-Cut and Diffusion

*Graph representation.* Let  $G = (V, E)$  be a weighted undirected graph, where the sets  $V$  are the graph nodes and  $E$  the edges connection nodes. In this paper, each node  $V_i$  represents a pixel  $i$  in an image  $I$  with support  $\Omega \subset \mathbb{R}^n$  where typically  $n = \{2, 3\}$ . The similarity between two pixels/nodes  $i$  and  $j$  in  $\Omega$  is measured by the edge function on the graph, namely  $w_{ij}$ . In the case of image segmentation, two pixels  $i$  and  $j$  that belong to the same object/class are said to be connected and define a measure  $w_{ij}$  close to unity. Inversely, two pixels  $i$  and  $j$  that do not belong to the same class are said to be not connected and define a measure  $w_{ij}$  close to zero. A standard construction approach for the weight matrix  $w_{ij}$  is as follows. Let  $h(i, j)$  be some general non-negative distance measure between nodes  $i$  and  $j$ , then the weight  $w_{ij}$  is computed with a Gaussian kernel of 0-mean and variance  $\sigma$  such that:

$$w_{ij} = \frac{1}{Z} \exp\left(-\frac{h(i, j)}{\sigma^2}\right),$$

where  $\sigma$  is the scaling parameter and  $Z$  is the normalization factor.

*Image Feature.* The distance  $h(i, j)$  depends on image feature. The choice of features is difficult and critical to get an optimal segmentation result. For piecewise smooth and constant images, the gray-level value can be enough. For texture images, a feature vector at each pixel computing from a filter bank (as suggested in e.g. [11]) can be efficient. A recent promising image feature to represent and process textures is the image intensity patch around the current pixel. The patch idea as feature vector was first introduced for texture synthesis [9, 12, 13] then for image denoising. Buades et al. in [6] proposed to compute the weight matrix with

patch differences and denoise the image with a non-local averaging. Gilboa-Osher in [7] and Szlam-Maggioni-Coifman in [14] proposed non-local/graph-based image denoising models based on patch differences. Finally, Bresson and Chan in [15] proposed a variational unsupervised segmentation method also based on patch differences. In this paper, we will use the graph of image patches of Buades et al.

*Min-Cut.* By definition a binary cut partitions a graph into two subsets. This partition process for graphs can be used for image segmentation when we want to find an object and the background. In optimization theory about maximum flows in flow networks [8], the optimal partition of the graph  $V$  into two sets  $A$  and  $B$  such that  $A \cup B = V$  and  $A \cap B = \emptyset$  can be computed by finding the minimal cut (min-cut), i.e. the minimization of the inter-similarity between two sets  $A$  and  $B$  of  $V$ . In other words, given two particular nodes  $s \in A$  and  $t \in B$  in the graph, the min-cut partition can be written as:

$$\min - \text{cut}(A, B) = \min_x \sum_{x_i > 0, x_j < 0} -w_{ij}x_i x_j, \quad (1)$$

where  $x$  is a  $N$ -dimensional indicator vector, with  $N = \text{card}(V)$ , such that  $x_i = 1$  if node  $i \in A$ , and  $x_i = -1$  otherwise. The min-cut partition gives the minimal capacity defined as the total weight between the nodes for each subset  $A$  and  $B$ . The min-cut approach has been applied to several computer vision problems, see [16, 17] for image restoration, [18, 19, 4, 5] for image segmentation, [20] for stereo and motion, and [21] for texture synthesis.

*Diffusion.* In the case of binary partition of a graph, min-cut partition and diffusion are equivalent. Indeed, let us denote  $W$  the symmetric matrix such that  $W(i, j) = w(i, j)$ . Then, the graph partition energy defined in (1) is equivalent to a Graph Laplacian operator:

$$\text{cut}(A, B) = \frac{1}{8} \sum w_{ij}(x_i - x_j)^2. \quad (2)$$

*Proof.* See e.g. [10]. Let  $\frac{x+1}{2}$  (resp.  $\frac{x-1}{2}$ ) be the indicator function for  $x_i > 0$  (resp.  $x_i < 0$ ). Then (1) can be written in the matrix form as follows:

$$\text{cut}(A, B) = \frac{1}{4} x^T (D - W)x,$$

where  $D$  is a  $N \times N$  diagonal matrix with  $d_i = \sum_j w(i, j)$  on its diagonal, the matrix  $D - W$  is called the Graph-Laplacian. It yields to:

$$x^T (D - W)x = \frac{1}{2} \sum_{i,j} w_{ij}(x_i - x_j)^2.$$

The weighted Graph Laplacian corresponds to a finite difference approximation of the continuous Laplacian operator. The graph Laplacian can also be a non-local operator.

*Semi-supervised segmentation.* We observe that min-cut partitioning algorithms are defined as semi-supervised segmentation techniques. The min-cut seeks for the optimal partition of the graph given particular nodes called the source "s" and the sink "t". Hence, it is easy to assign some pixels as source and some as sink if the pixel belongs to the object or the background. Several graph-based partitioning methods have been proposed in the literature s.a. [18, 22–24, 19]. The previous papers are based on discrete minimization methods to compute the min-cut given the labels. In this paper, we propose a continuous minimization method to solve the min-cut problem with labels and non-local image information.

### 3 Proposed Segmentation Method

#### 3.1 Continuous Min-Cut

*Energy minimization problem.* In this section a new non-local semi-supervised segmentation algorithm is introduced. The algorithm relies on the continuous formulation of the discrete min-cut problem defined as:

$$\begin{aligned} \min - \text{cut}(A, B) &= \min_x \sum_{i,j} w_{ij} (x_i - x_j)^2 & (3) \\ \text{s.t.} \quad x_k &= +1, \quad \forall k \in S \\ x_k &= -1, \quad \forall k \in T, \end{aligned}$$

where  $S$  are the labels selected for the object and  $T$  are the labels assigned to the background.

We propose the continuous min-cut (CMC) problem as follows (which is a constrained minimization problem w.r.t. a real-valued function  $u$ ):

$$\begin{aligned} CMC(u) &= \min_u \frac{1}{2} \iint_{\Omega \times \Omega} w(x, y) (u(x) - u(y))^2 dx dy \\ \text{s.t.} \quad u(x) &= 1, \quad \forall x \in S \\ u(x) &= 0, \quad \forall x \in T, \end{aligned}$$

which is equivalent to this unconstrained minimization problem for  $u$ :

$$\begin{aligned} CMC(u) &= \min_u \frac{1}{2} \iint_{\Omega \times \Omega} w(x, y) (u(x) - u(y))^2 dx dy + \int_{\Omega} \lambda(x) (u - u_0)^2 dx, \quad (4) \\ \text{where} \quad u_0(x) &= \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \in T \end{cases} \quad \text{and} \quad \lambda(x) = \begin{cases} \infty & \text{if } x \in S \cup T \\ 0 & \text{otherwise} \end{cases}, \end{aligned}$$

where function  $\lambda$  provides the degree of confidence with respect to the labels.

*Non-local  $H^1$  energy.* The first term of (4) is deduced from (3) using the change of variable  $u_i = \frac{x_i + 1}{2} \in \{0, 1\}$  then relaxing  $u_i$  to  $[0, 1]$ . This term is also known

as the non-local  $H^1$  energy ([25, 7]) defined as:

$$H_G^1(u) = \frac{1}{2} \iint_{\Omega \times \Omega} w(x, y)(u(x) - u(y))^2 dx dy = \frac{1}{2} \int_{\Omega} |\nabla_G u|^2 dx = \|u\|_{H_G^1}, \quad (5)$$

where  $|\nabla_G u|^2 := \int_{\Omega} w(x, y)(u(x) - u(y))^2 dy$  is the square norm of the continuous graph gradient of  $u$ . The optimality condition for (5) is:

$$\int_{\Omega} w(x, y)(u(x) - u(y)) dy = \Delta_G u = 0,$$

where  $\Delta_G u$  is the continuous graph Laplacian of  $u$ .

*Labels.* The second term of (4) introduces the hard constraint of labels in the energy minimization approach. This term comes from Unger et al. in [26], which incorporates seed points (assigned either to the object or to the background) in the geodesic active contour/snake model [27]. This term constrains function  $u$  to be equal to  $u_0$  for  $x \in S \cup T$  and being equal to anything else for  $x \notin S \cup T$ . Function  $\lambda$  is highly discontinuous, which requires some regularization process to handle it. Unger et al. proposed a splitting operation to solve this problem. A new function  $v$  is introduced s.t.:

$$\min_u \int_{\Omega} \lambda(x)(v - u_0)^2 dx + \frac{1}{2\theta} \|u - v\|_2^2,$$

where the term  $\|u - v\|_2^2$  forces  $v \approx u$  as  $\theta \rightarrow 0$ . The optimality solution w.r.t.  $v$  leads to:

$$v = \frac{2\lambda(x)\theta u_0 - u}{2\lambda(x)\theta - 1} = \begin{cases} u & \text{if } \lambda \rightarrow 0 \\ u_0 & \text{if } \lambda \rightarrow \infty \end{cases}.$$

### 3.2 Proposed Semi-Supervised Segmentation Algorithm: Continuous Min-Cut + TV

*Final model.* The previous section introduced the continuous formulation of the min cut problem. In this section, we proposed to merge the continuous min-cut with the Total Variation (TV) energy. The TV term offers two advantages. First, TV regularizes the geometry of the contour between classes (object and background). Experiments showed that the continuous min-cut can provide irregularities along the contour. Second, Shi and Malik in [10] observed that the min-cut algorithm tends to favor misclassification of small sets, which are smoothed out with the TV regularization process.

Finally, we propose the following energy minimization model for semi-supervised segmentation:

$$E(u) = \|u\|_{H_G^1} + \int_{\Omega} \lambda(x)(u - u_0)^2 dx + \beta \|u\|_{TV}, \quad (6)$$

where  $\|u\|_{TV} = \int_{\Omega} |\nabla u| dx$ .

*Minimization process.* A direct use of the calculus of variation to (6) will produce a very slow minimization process. We propose to use a splitting operation to minimize  $E$  more efficiently. We introduce two new functions  $v, s$  s.t.:

$$E(u, v, s) = \|u\|_{H_G^1} + \int_{\Omega} \lambda(x)(v - u_0)^2 dx + \beta \|s\|_{TV} + \frac{1}{2\theta_v} \|u - v\|_2^2 + \frac{1}{2\theta_s} \|s - v\|_2^2. \quad (7)$$

Then,  $v, s$  being fixed, we search for  $u$  as the solution of  $\min_u \|u\|_{H_G^1} + \frac{1}{2\theta_v} \|u - v\|_2^2$ , which is given by a fixed point method as  $u = \frac{\theta_v \int_{\Omega} w(x,y)u(y)dy + v(x)}{\theta_v \int_{\Omega} w(x,y)dy + 1}$ . Functions  $u, s$  being fixed, we search for  $v$  as the solution of  $\min_v \int_{\Omega} \lambda(x)(v - u_0)^2 dx + \frac{1}{2\theta_v} \|u - v\|_2^2 + \frac{1}{2\theta_s} \|s - v\|_2^2$ , which is given by  $v = \frac{\theta_s v + \theta_v s}{\theta_s + \theta_v}$  if  $\lambda = 0$  and  $v = u_0$  if  $\lambda = \infty$ . Functions  $u, v$  being fixed, we search for  $s$  as the solution of  $\min_s \beta \|s\|_{TV} + \frac{1}{2\theta_s} \|s - v\|_2^2$ , which solution is given e.g. the Projection algorithm of Chambolle [28].

We propose the following iterative scheme for minimizing energy (7):

$$\begin{cases} u^{n+1} = \frac{\theta_v \int_{\Omega} w(x,y)u^n(y)dy + v^n(x)}{\theta_v \int_{\Omega} w(x,y)dy + 1} \\ v^{n+1} = \begin{cases} \frac{\theta_s u^{n+1} + \theta_v s^n}{\theta_s + \theta_v} & \text{if } \lambda = 0 \\ u_0 & \text{if } \lambda = \infty \end{cases}, \quad n \geq 0 \\ s^{n+1} = v^{n+1} - \theta \operatorname{div} p^n \\ p^{n+1} = \frac{\beta(p^n + 1/8 \nabla(\operatorname{div} p^n - v^{n+1}/\theta_s))}{\beta + 1/8 |\nabla(\operatorname{div} p^n - v^{n+1}/\theta_s)|} \end{cases} \quad (8)$$

### 3.3 Some Properties of the Models (6) and (7)

*Convexity.* Both energy minimization models (6) and (7) are strictly convex (since the  $H_G^1$  term is strictly convex), which implies the existence of a unique minimizing solution independently of the initial condition. Hence, even using gradient descent approaches, the algorithm does not get stuck in a local minimum. Thus, as far as the labels are correctly defined, the results will be independent of the initialization.

*Relation with the original min-cut problem (1) or equivalently (3).* The continuous min-cut problem (4) has the same solution as the discrete min-cut problem when considering characteristic/indicator functions of sets, i.e.:

$$\min_{1_A} \{E_{CMC}(u = 1_A)\} = \min_A \{cut(A, \Omega \setminus A)\}. \quad (9)$$

We remind that we relax function  $u$  to be between  $[0, 1]$  to define a continuous version of the min-cut algorithm, which can be minimized with continuous

minimization tools. Then, the segmentation result is given by thresholding the minimizer  $u$  of (6) with any value in  $(0, 1)$ . Unfortunately, we do not have a theorem like in [29, 30] to state that the optimal solution can be thresholded at any value in  $(0, 1)$  and yields to the same result. However, experimental results suggest that minimizing solutions of (1) are close to binary functions, and so provide the same result for any threshold value in  $(0, 1)$ .

*Non-trivial steady state solution of (6).* The final steady state solution of (6) is not the mean value of the initial function. Call  $\bar{u}_{t=0} = \int u_{t=0}$  the mean value function. It is easy to show by contraction that  $\bar{u}_{t=0}$  is not solution to (6). If  $u_{t=\infty} = \bar{u}_{t=0}$  then  $E(u_{t=\infty}) = \int_{\Omega} \lambda(x)(\bar{u}_{t=0} - u_0)^2 dx > 0$ , and the minimizer is thus given by  $\bar{u}_{t=0} = u_0$ . However,  $u_0(x) = 1 \ \forall x \in S$ ,  $0 \ \forall x \in T$ . Thus,  $\bar{u}_{t=0} \neq u_0$ .

We notice that Gilboa and Osher in [7] also use the energy  $\|u\|_{H_G^1}$  to perform semi-supervised segmentation. However, they did not use a term to constraint the labels as in this work. They minimized energy  $\|u\|_{H_G^1}$  starting with a trinary initial function  $u_{t=0} = \{-1, 0, 1\}$  (labelled pixels for the object are assigned to the value 1 and those for the background to the value  $-1$ ). However, the minimizing solution is the mean value function  $u_{t=\infty} = \bar{u}_{t=0}$ . Hence, this algorithm requires to stop the diffusion process.

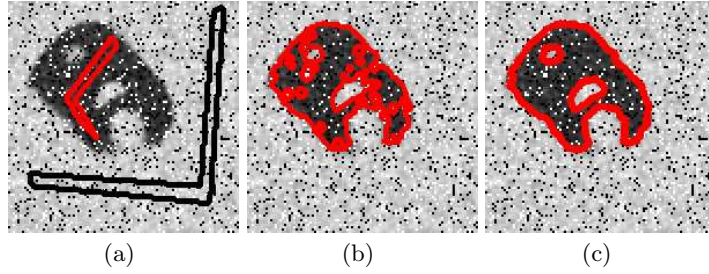
## 4 Results

This section presents some results of the proposed semi-supervised segmentation algorithm. The graph is defined from local and non-local image information:

$$w(i, j) = \begin{cases} \frac{|i-j|_2}{\sigma_1^2} + \frac{|F(i)-F(j)|_2}{\sigma_2^2} & \text{if } i, j \in \mathcal{N}_{a \times a}(i) \\ 0 & \text{otherwise} \end{cases}, \quad (10)$$

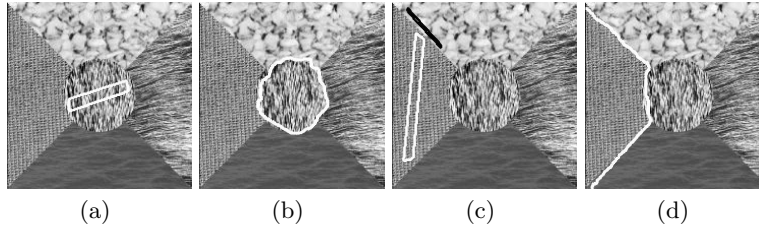
where  $\mathcal{N}_{a \times a}(i)$  is a square window of size  $a \times a$  around  $i$ . The computational cost of the similarity between pixel on the whole image is very expensive, however we chose to simply select points in a close neighborhood. This implies the supposition that if two points are far away, they are not connected. From  $a^2$  neighbors only the  $cl = 8$  closest points are selected. The feature vector  $F$  is a square patch of size  $f \times f$  centered on each pixel. The segmentation is driven by (8). The initial condition for  $u$ ,  $v$  and  $s$  are given by the label  $S$ , i.e.  $u = v = s = 1$  if  $x \in S$  and  $u = v = s = 0$  otherwise. With an unoptimized Matlab implementation, the graph computation lasts approximatively 15 seconds and the segmentation is performed in approximatively 1 minute. The image size is  $128 \times 128$ .

*TV Regularization Effect* The importance of the TV-Regularization effect is emphasized in this paragraph. A salt-and-pepper noise is added on a two-phase image with different means 1(a). The inside and outside labels are presented on Figure 1(a). The results show that if the TV regularization is not performed then the segmentation fails (Fig 1(b)). When the TV regularization is used, then the segmentation succeeds.



**Fig. 1.** Application of our algorithm on a image with a salt-and-pepper noise. (a)Initialization (b)The segmentation result without TV-regularization.(c)The segmentation result with TV-regularization.

*Texture Images* We apply our algorithm to a synthetic texture image composed of five different patterns. Figures 2(a) and 2(c) show the initializations and Figures 2(b) and 2(d) the corresponded results. The patch size is chosen to be  $9 \times 9$  which correspond to the pattern size for the two selected textures.



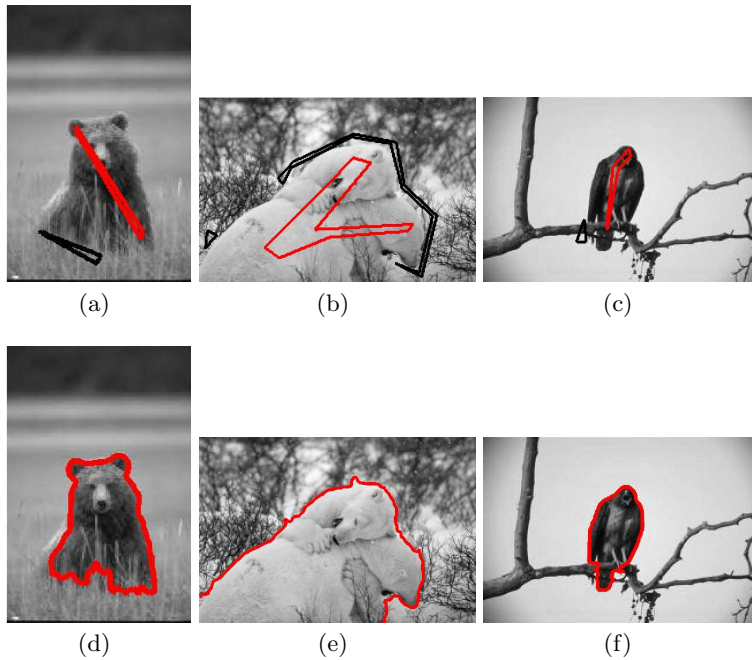
**Fig. 2.** Results on synthetic textures. (a) and (c) Initializations. (b) and (d) results.

*Natural Images* We apply now our algorithm to a set of natural images taken from the Berkeley segmentation dataset [31]. In the first column of Figure 3, the inside and outside labels are shown and in the second column the segmentation results.

*Color Images* We consider the simple case of Red-Green-Blue (RGB) channels. The first step consists of computing the graph by taking into account each channel, i.e.  $F = (F_r, F_g, F_b)$ , where respectively the red, green and blue feature channel. Images are also taken from the Berkeley segmentation dataset [31]. In the first column of Figure 3, the inside and outside labels are shown and in the second column the segmentation results.

*Medical Images* We apply our segmentation algorithm on 2-D medical images of CT scans of the abdomen and the head and neck. Figures 5(a) and 5(c) present





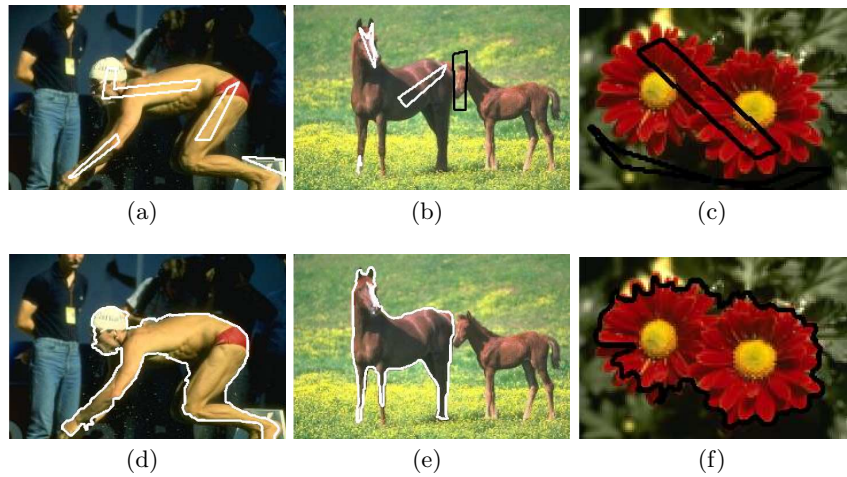
**Fig. 3.** Results on real-world images from the Berkeley dataset. Left column: Initial labels. Right column: Segmentation Result

the inside and outside initial labels. Figures 5(b) and 5(d) show the segmentation results. For the liver segmentation, the label on the background (black) prevents the diffusion from capturing as well the heart. The segmentation of the structures in the neck are challenging and the result that we obtain are promising.

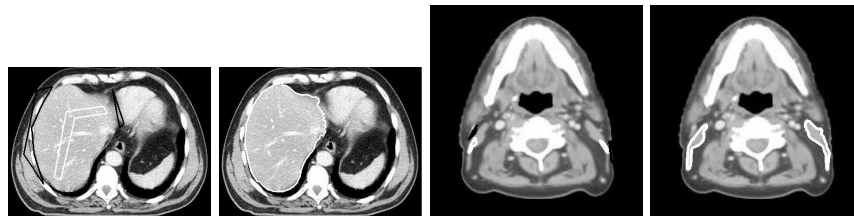
## 5 Discussion and Conclusion

In this paper, a non-local semi-supervised segmentation method has been proposed. The success of graph partitioning algorithms for image segmentation has motivated this work. Our objective was to translate the discrete min-cut algorithm into a non-local continuous min-cut algorithm. The addition of hard constraints with the source and sink labels has been done naturally in the proposed continuous framework. Besides, it has also been easy to introduce new terms such as the TV term that regularizes the geometry of the boundary between the object and the background. The non-local continuous min-cut is also equivalent to a diffusion process. The diffusion is done on the graph of image intensity patches, which holds semi-local and non-local image information useful to segment textures and complex patterns. Our semi-supervised segmentation has provided promising segmentation results for textures and real-world objects.

Future work will focus on comparing the efficiency of our segmentation algorithm with other related semi-supervised segmentation algorithms. We would like also to extend our method to 3-D medical images.



**Fig. 4.** Results on real world color images from the Berkeley dataset. Left column: Initial labels. Right column: Segmentation Result



**Fig. 5.** First row, Segmentation of the liver. (a)Initial labels. (b)Segmentation Result. Second row, Segmentation of the lateral muscles on the neck. (c)Initial labels. (d)Segmentation Result. (e)Zoom on the segmentation of the muscles

## 6 Acknowledgements

Nawal Houhou was supported by Swiss National Science Foundation #205320-101621, Xavier Bresson was supported by ONR N00014-03-1-0071 and ONR MURI subcontract from Stanford University and Arthur Szlam was supported

by NSF DMS-0811203. The authors would like also to thank the referees for their constructive comments.

## References

1. Strang, G.: Maximal Flow Through A Domain. *Mathematical Programming* **26(2)** (1983) 123–143
2. Protiere, A., Sapiro, G.: Interactive image segmentation via adaptive weighted distances. *IEEE Transactions on Image Processing* **16(4)** (2007) 1046–1057
3. Cremers, D., Fluck, O., Rousson, M., Aharon, S.: A Probabilistic Level Set Formulation for Interactive Organ Segmentation. In: *SPIE*. (2007)
4. Appleton, B., Talbot, H.: Globally minimal surfaces by continuous maximal flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28(1)** (2006) 106–118
5. Unger, M., Pock, T., Cremers, D., Bischof, H.: Tvseg - interactive total variation based image segmentation. In: *British Machine Vision Conference (BMVC)*, Leeds, UK (September 2008)
6. Buades, A., Coll, B., Morel, J.: A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation* **4(2)** (2005) 490–530
7. Gilboa, G., Osher, S.: Nonlocal linear image regularization and supervised segmentation. *Multiscale Modeling and Simulation* **6(2)** (2007) 595–630
8. P. Elias, A.F., Shannon, C.E.: Note on Maximum Flow Through a Network. *IRE Transactions on Information Theory* **2** (1956) 117119
9. Efros, A., Leung, T.: Texture Synthesis by Non-Parametric Sampling. *IEEE International Conference on Computer Vision* **2** (1999) 10–33
10. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** (2000) 888–905
11. Malik, J., Belongie, S., Leung, T., Shi, J.: Contour and texture analysis for image segmentation. *International Journal of Computer Vision* **43(1)** (2001) 7–27
12. Efros, A., Freeman, W.: Image quilting for texture synthesis and transfer. In: *Proceedings of the Conference on Computer graphics and interactive techniques, SIGGRAPH*, New York, NY, USA, ACM (2001) 341–346
13. Liang, L., Liu, C., Xu, Y., Guo, B., Shum, H.: Real-time texture synthesis by patch-based sampling. *ACM Trans. Graph.* **20(3)** (2001) 127–150
14. Szlám, A., Maggioni, M., Coifman, R.: Regularization on Graphs with Function Adapted Diffusion Processes. *Journal of Machine Learning Research* (2008)
15. Bresson, X., Chan, T.: Non-local Unsupervised Variational Image Segmentation Models, *UCLA CAM Report 08-67* (2008)
16. Wu, Z., Leahy, R.: An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15(11)** (1993) 1101–1113
17. Ishikawa, H., Geiger, D.: Segmentation by grouping junctions. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (1998) 125–131
18. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. Volume 1. (2001) 105–112
19. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient n-d image segmentation. *Int. J. Comput. Vision* **70(2)** (2006) 109–131

20. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23** (2001) 2001
21. Kwatra, V., Schdl, A., Essa, I., Turk, G., Bobick, A.: Graphcut textures: Image and video synthesis using graph cuts. *Proceedings of the Conference on Computer graphics and interactive techniques, SIGGRAPH* **22**(3) (July 2003) 277–286
22. Blum, A., Chawla, S.: Learning from labeled and unlabeled data using graph mincuts. In: *Proceedings of the Eighteenth International Conference on Machine Learning*, San Francisco, CA, USA, Morgan Kaufmann Publishers Inc. (2001) 19–26
23. Yu, S., Shi, J.: Segmentation given partial grouping constraints. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(2) (2004) 173–183
24. Grady, L., Funka-lea, G.: Multi-label image segmentation for medical applications based on graph-theoretic electrical potentials. In: *Proceedings of the European Conference on Computer Vision*, Springer (2004) 230–245
25. Zhou, D., Scholkopf, B.: A Regularization Framework for Learning from Graph Data. In: *Workshop on Statistical Relational Learning and Its Connections to Other Fields*. (2004)
26. Unger, M., Pock, T., Bischof, H.: Continuous globally optimal image segmentation with local constraints. In: *Computer Vision Winter Workshop 2008*. (2008)
27. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic Active Contours. *International Journal of Computer Vision* **22**(1) (1997) 61–79
28. Chambolle, A.: An Algorithm for Total Variation Minimization and Applications. *Journal of Mathematical Imaging and Vision* **20**(1-2) (2004) 89–97
29. Chan, T., Esedoglu, S., Nikolova, M.: Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models. *SIAM Journal on Applied Mathematics* **66**(5) (2006) 1632–1648
30. Bresson, X., Esedoglu, S., Vandergheynst, P., Thiran, J., Osher, S.: Fast Global Minimization of the Active Contour/Snake Models. *Journal of Mathematical Imaging and Vision* **28**(2) (2007) 151–167
31. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Volume 2. (July 2001) 416–423